

METHOD AND APPARATUS FOR ESTIMATION OF COMPONENT GATING AND SHORTAGE RISK IN MANUFACTURING OPERATIONS

Balazs Kralik, Michael Goldbach, and Paul Dagum

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of the invention claims priority from a provisional application entitled "METHOD AND BUSINESS PROCESS FOR ESTIMATION OF COMPONENT GATING AND SHORTAGE RISKS IN ASSEMBLE-TO-ORDER MANUFACTURING OPERATIONS" (U.S. Provisional Application Number 60/229,840), filed August 31, 2000, having Balazs Kralik, Michael Goldbach, and Paul Dagum as inventors, and having as assignee Rapt, Inc., the assignee of the present invention. The aforementioned provisional patent application is incorporated herein by reference, in its entirety and for all purposes.

[0002] This invention is also related to U.S. Provisional Patent Application No. 60/213,189, filed June 21, 2000, which is incorporated by reference herein. It is also related to U.S. Patent Application Nos. 09/412,560, filed October 5, 1999, and 09/491,461, filed January 26, 2000, both of which are incorporated by reference herein.

BACKGROUND OF THE INVENTION

Field of the Invention

[0003] This invention relates generally to manufacturing resource allocation, and more particularly, to risk management in the procurement of unfinished goods by estimation of component gating and shortage risk.

Description of the Related Art

[0004] An assemble-to-order (ATO) manufacturing process is one in which products are manufactured from raw components only as orders are received. While ATO processes are more efficient than manufacture-to-stock (MTS) operations, in which products are produced before demand is known, they are not without significant risks. If component availability is below what is required to meet product demand, then fulfillment of orders is either delayed or

fails. Conversely, if excess component inventory remains after demand is met, the excess is either returned to the supplier, usually at a loss, or held until future need. Thus any component inventory that does not exactly meet demand incurs some financial loss for the manufacturer. Being able to predict the risk of component shortage for each component allows a manufacturer to make more informed business decisions.

SUMMARY OF THE INVENTION

[0005] In one embodiment, a method for estimation of component gating risk in manufacturing operations is disclosed. The method includes generating an altered component plan by altering a component plan for a component, computing a mean production value using the altered component plan, and computing the component gating risk using the mean production value.

[0006] In another embodiment, a method for estimation of component gating risk in manufacturing operations is disclosed. The method includes generating an increased component plan by increasing a component plan for a component, computing a first mean production value using the increased component plan, generating a decreased component plan by decreasing the component plan for the component, computing a second mean production value using the decreased component plan, and computing the component gating risk using the first and the second mean production values.

[0007] In still another embodiment, a method for estimation of component shortage risk in manufacturing operations is disclosed. The method includes identifying a component and computing r . In this embodiment, r is the component shortage risk, and is defined by

$$r = \int_0^d N(x, a \bullet \mu, \sqrt{a^T \Sigma a}) dx$$

where d is a maximum expeditable level of the component, $N(x, \mu, \sigma)$ is a normal density function having a mean equal to μ and a variance equal to σ , a is a vector of connect rates for the component, μ is a mean demand, and Σ represents demand covariance.

[0008] The foregoing is a summary and thus contains, by necessity, simplifications, generalizations and omissions of detail; consequently, those skilled in the art will appreciate

that the summary is illustrative only and is not intended to be in any way limiting. As will also be apparent to one of skill in the art, the operations disclosed herein may be implemented in a number of ways, and such changes and modifications may be made without departing from this invention and its broader aspects. Other aspects, inventive features, and advantages of the present invention, as defined solely by the claims, will become apparent in the non-limiting detailed description set forth below.

BRIEF DESCRIPTION OF THE DRAWINGS

[0009] The present invention may be better understood, and its numerous objects, features, and advantages made apparent to those skilled in the art by referencing the accompanying drawings.

[0010] FIG. 1 is a schematic diagram illustrating a "p2c2" relationship between products and components.

[0011] FIG. 2 is a graph of demand showing regions of shortage for components from FIG. 1.

[0012] FIG. 3 is a graph of demand showing regions of gating for components from FIG. 1.

[0013] FIG. 4 is a flow diagram illustrating one embodiment of a procedure for calculating gating risk.

[0014] FIG. 5 is a flow diagram illustrating one embodiment of a procedure for calculating shortage risk.

[0015] ~~Fig. 1~~ FIG. 6 is a block diagram illustrating a network environment in which ~~commercial transaction~~ processing according to embodiments of the present invention may be practiced.

[0016] ~~Fig. 2~~ FIG. 7 is a block diagram illustrating a computer system suitable for implementing embodiments of the present invention.

[0017] ~~Fig. 3~~ FIG. 8 is a block diagram illustrating the interconnection of the computer system of ~~Fig. 2~~ FIG. 7 to client and host systems.

[0018] The use of the same reference symbols in different drawings indicates similar or identical items.

DETAILED DESCRIPTION OF THE INVENTION

DETAILED DESCRIPTION

[0019] The following is intended to provide a detailed description of examples of an ~~example of~~ the invention and should not be taken to be limiting of the invention itself. Rather, any number of variations may fall within the scope of the invention which is defined in the claims following the description.

Introduction

[0020] Described herein are methods and systems that may be usable for ~~The present invention addresses this need by providing a novel method for~~ estimating component shortage risk and component gating risk. A gating component is one that is the most short and therefore determines the allowable level of production. Gating risk for a particular component is the risk that the component will be the gating component. Estimation of component shortage and gating risks is particularly complicated when a large number of products are made from a large number of the same components, and thus the techniques described herein may be ~~present invention is particularly~~ advantageous under such conditions.

[0021] When possible, manufacturing operations may find it beneficial to produce their outputs in an ATO fashion. While MTS operations produce outputs before demand for them is revealed, ATO operations generally produce outputs only as orders arrive.

[0022] In an ATO operation, the availability of components is a significant variable that controls ability to produce. If component availability is limited (e.g., because inventory is limited, or because supplier commitments to ship the component on demand are insufficient), then fulfillment of demand either fails or is delayed.

[0023] If component planners must negotiate component supply contracts (which specify the number of each component to ship in a planning period) before demand is realized, there is a residual risk that some components will be in short supply. Estimating this risk may help inform supply contract negotiation.

[0024] Component Shortage Risk quantifies the risk that in a planning period there will be more need than availability for a particular component. In other words, Component Shortage Risk for a particular component is the risk that there will be a shortage of that particular component.

[0025] Component Gating Risk quantifies the risk that a particular component will be the cause of a failure to fulfill an order. In many situations more than one component may be short. But one of the short components—the one that is the most short—is the “gating component.” That is, the gating component is the component that determines the level of production. Component Gating Risk for a particular component is the risk that that particular component will be the gating component.

[0026] FIG. 1 is a schematic diagram illustrating a “p2c2” model 10 of products and components. In this model, two components 15 and 16 are used to make two products 13 and 14. Component 1 (15) is associated with Product 1 (13), as indicated by a line connecting these blocks. Similarly, Component 2 (16) is associated with Product 2 (14). An additional association is also indicated between Component 2 and Product 1.

[0027] The behavior of model 10 is further described in the graph of FIGS. 2 and 3. These graphs illustrate a behavior that is consistent with the associations shown in FIG. 1, where each unit of Product 1 is shown as depending on a unit of Component 1 and a unit of Component 2, and where each unit of Product 2 is shown as depending on a unit of Component 2.

[0028] FIG. 2 is a graph 20 of demand showing regions of shortage for components in the model 10. The horizontal axis represents the demand for Product 1, denoted as x_1 . The vertical axis represents the demand for Product 2, denoted as x_2 . FIG. 2 further indicates the quantities d_1 and d_2 , which are the maximum component availabilities of Component 1 and Component 2, respectively.

[0029] The graph 20 includes four regions 23, 24, 25, and 26, which are separated by a vertical line 21 and a sloped line 22. The vertical line 21 represents the situation where the demand for Product 1 is equal to the maximum component availability of Component 1, i.e.: $x_1 = d_1$. The sloped line 22 represents the situation where the demand for Product 1 plus the demand for Product 2 is equal to the maximum component availability of Component 2, i.e.: $x_1 + x_2 = d_2$.

[0030] Region 23 is located to the left of the vertical line 21 and below the sloped line 22. Thus, region 23 has values of the demands x_1 and x_2 that are feasible: in this region there is enough of Component 1 and of Component 2 to meet the demands for Product 1 and Product 2. This may be seen from the fact that in region 23, $x_1 \leq d_1$ and $x_1 + x_2 \leq d_2$. Thus, in this region 23 the demanded number of units (x_1) of Product 1 does not exceed the available supply (d_1) of Component 1, which is needed for Product 1. Also, in this region 23 the demanded number of units of Product 1 combined with the demanded number of units of Product 2 ($x_1 + x_2$) does not exceed the available supply (d_2) of Component 2, which is needed booth for Product 1 and for Product 2.

[0031] Region 24 is located to the right of the vertical line 21 and below the sloped line 22. Thus, region 24 has values of the demands x_1 and x_2 that would result in a shortage of Component 1: in this region there is enough of Component 2 but not enough of Component 1 to meet the demands for Product 1 and Product 2. This may be seen from the fact that in region 24, $x_1 + x_2 \leq d_2$, but $x_1 > d_1$. Thus, in this region 24 the demanded number of units (x_1) of Product 1 exceeds the available supply (d_1) of Component 1, which is needed for Product 1. The demands in region 24 would thus result in a shortage of Component 1.

[0032] Region 25 is located to the left of the vertical line 21 and above the sloped line 22. Thus, region 25 has values of the demands x_1 and x_2 that would result in a shortage of Component 2: in this region there is enough of Component 1 but not enough of Component 2 to meet the demands for Product 1 and Product 2. This may be seen from the fact that in region 25, $x_1 \leq d_1$, but $x_1 + x_2 > d_2$. Thus, in this region 25 the demanded number of units (x_1) of Product 1 can be met by the available supply (d_1) of Component 1, which is the only component needed for Product 1. However, the combined demand ($x_1 + x_2$) for Products 1 and 2 can not be met by the available supply (d_2) of Component 2, one unit of which is

needed for each of Products 1 and 2. The demands in region 25 would thus result in a shortage of Component 2.

[0033] Region 26 is located to the right of the vertical line 21 and above the sloped line 22. Thus, region 26 has values of the demands x_1 and x_2 that would result in a shortage of both Component 1 and Component 2: in this region there is not enough of Component 1 and not enough of Component 2 to meet the demands for Product 1 and Product 2. This may be seen from the fact that in region 26, $x_1 > d_1$ and $x_1 + x_2 > d_2$. Thus, in this region 26 the demanded number of units (x_1) of Product 1 can not be met by the available supply (d_1) of Component 1, which is the only component needed for Product 1. Further, the combined demand ($x_1 + x_2$) for Products 1 and 2 can not be met by the available supply (d_2) of Component 2, one unit of which is needed for each of Products 1 and 2. The demands in region 26 would thus result in a shortage of Component1 and a shortage of Component 2.

[0034] Production is feasible if the demands x_1 and x_2 fall in the feasible region 23. In those situations, the demands x_1 and x_2 can be met by the available supply of components d_1 and d_2 . The feasible region may be denoted by the symbol Ω . In the model 10 depicted in FIGS. 1 and 2, the feasible region Ω is defined by the boundaries $x_1 \leq d_1$ and $x_1 + x_2 \leq d_2$. It may be seen that these conditions can be written more generally as:

$$\Omega = \{\bar{a}_i \bullet \bar{x} \leq d_i \quad \forall i\} \quad (1).$$

In Eq. 1, vector \bar{x} in Eq. 1 represents the demand for products. The quantity d_i in Eq. 1 represents the maximum component availability for component i . The vector \bar{a}_i are the "connect rates" for component i . Connect rates between products and components represent the number of components required for the manufacture of one unit of product. Connect rates may be used to describe the consumption of components for each unit of the various products.

[0035] In the two-product model 10 discussed above, \bar{x} could be written as (x_1, x_2) . The quantity d_i would be d_1 for Component 1 ($i=1$), and d_2 for Component 2 ($i=2$).

[0036] A shortage event occurs if the realized demand \bar{x} is such that $\bar{a}_i \bullet \bar{x} \leq d_i$. With reference to FIGS. 1 and 2, it may be seen that the connect rates \bar{a}_i represent the relationship between products and components. In the model 10 discussed above, it may be readily seen that the connect rates \bar{a}_1 for Component 1 have a value of (1,0), since one unit of Component 1 is used for each unit of Product 1, and zero units of Component 1 are used for each unit of Product 2. Similarly, the connect rates \bar{a}_2 for Component 2 would have a value of (1,1), since in the model 10 one unit of Component 2 is used for each unit of Product 1, and one unit of Component 2 is used for each unit of Product 2. With these values of \bar{a}_1 and \bar{a}_2 for the model 10, it may be readily seen that Eq. 1 directly indicates that the feasible region Ω has demand values (x_1 and x_2) for which $x_1 \leq d_1$ and $x_1 + x_2 \leq d_2$. This result from Eq. 1 matches the graphical analysis discussed above for the feasible region 23.

[0037] A forecasted set of demands may be represented by the vector variable $\bar{x} = (x_1, x_2, \dots)$, which includes a scalar value for each of the products in a production scheme. This demand vector may be understood as having an expected value that is represented by $\bar{\mu} = (\mu_1, \mu_2, \dots)$ and a probability distribution that allows for the occurrence of other values of the demand vector. The probability distribution of the demand vector \bar{x} may be modeled as a normal distribution with a mean value of $\bar{\mu}$ and a distribution that is represented by a covariance matrix denoted as Σ .

[0038] The vector notation may also be generalized to describe a general number of components. For example, the vector quantity $\bar{d} = (d_1, d_2, \dots)$ may be used to indicate the component availability. This vector includes a scalar value for each of the components in a production scheme.

[0039] As illustrated in FIG. 2, the expected value $\bar{\mu}$ 28 of the demand vector may be located in the feasible region 23. However, as illustrated in the figure, the distribution 29 of the possible demand may extend into one or more of the non-feasible regions 24, 25, and 26. The distribution 29 is represented schematically by an ellipse. This ellipse 29 indicates an iso-probability curve that connects values of \bar{x} that are equally likely. Because the

distribution has some overlap with the non-feasible regions, there is a non-zero chance that the demand for a set of products may not be feasible with the component availability \bar{d} . There may therefore be a risk that the actual demand vector will fall outside the feasible region, resulting in a shortage of one or more of the components. The shortage risk r_i for a component i may be understood as the probability that the demand for that component will exceed the maximum availability d_i of that component. This quantity may be calculated according to the following formula.

$$r_i = \int_{d_i}^{\infty} N(x, \bar{a}_i \bullet \bar{\mu}, \sqrt{\bar{a}_i^T \Sigma \bar{a}_i}) dx \quad (2)$$

[0040] In Eq. 2, it may be seen that x is an integration variable that runs through the non-feasible values of component i : from d_i to infinity. The function $N(x, \mu, \sigma)$ is the normal density distribution of a variable x with mean of μ and a variance of σ^2 . The integral in Eq. 2 thus evaluates the probability that a normally-distributed demand will require an amount of component i in excess of the available quantity d_i .

[0041] FIG. 3 is a graph of demand showing regions of gating for components from FIG. 1. Again, the horizontal axis represents the demand x_1 for Product 1 and the vertical axis represents the demand x_2 for Product 2. FIG. 3 also includes the vertical line 21 where $x_1 = d_1$, and the sloped line 22 where $x_1 + x_2 = d_2$.

[0042] FIG. 3 additionally includes a sloped line 38 that has one end at the intersection of lines 21 and 22, and which divides the combined shortage regions (24, 25, and 26 from FIG. 2) into two new regions 34 and 35. (The sloped line 38 may be defined by a policy vector, shown in bold, that represents a production policy for the various products and components.) Region 34 is to the left of the sloped line 38. Region 35 is to the right of the sloped line 38.

[0043] Region 35 represents situations in which Component 1 would completely run out if an attempt were made to produce the demanded amounts x_1 and x_2 of Products 1 and 2. That is, Component 1 is the gating component in region 35. Similarly, region 34 represents situations in which Component 2 would completely run out if an attempt were made to produce the demanded amounts x_1 and x_2 of Products 1 and 2. That is, Component 2 is the

gating component in region 34. Since the demand is met by the components in region 23, neither Component 1 nor Component 2 is gating in region 23.

[0044] The regions of gating risk 34 and 35 are illustrated in Figure 3. Intuitively, the component i is the gating component in situations where the demand is \bar{x} and the component availability is \bar{d} if the production $\bar{q}(\bar{x}, \bar{d})$ falls on the d_i -face of the feasible region Ω . (The complement Ω is the infeasible region, denoted as $\bar{\Omega}$.) More strictly, given a demand event \bar{x} , a component plan \bar{d} , and a production policy $\bar{q} : \mathfrak{R}^n, \bar{d} \in \mathfrak{R}^m \rightarrow \mathfrak{R}^n$, a component i may be understood as a gating component if and only if $\bar{q}(\bar{x}, \bar{d})$ falls on the face of the feasible region Ω associated with component i . In other words, a component i may be understood as a gating component if and only if:

$$\bar{a}_i \bullet \bar{q}(\bar{x}, \bar{d}) = d_i \quad (3)$$

[0045] For a production policy such as depicted in FIG. 3, it may be readily seen that the space where more than one component is gating is of measure zero, except in the degenerate case where there are two components 1 and 2 such that $\bar{a}_1 = \bar{a}_2$ and $d_1 = d_2$. This degenerate case of measure zero may be excluded from the following discussion.

[0046] Gating risk may also be understood as: given a demand \bar{x} , whichever component actually runs out completely upon producing the production amount $\bar{q}(\bar{x}, \bar{d})$ is the gating component.

[0047] Now consider the function $g_i(\bar{x}, \bar{d})$, where:

$$g_i(\bar{x}, \bar{d}) = \bar{a}_i \bullet \bar{q}(\bar{x}, \bar{d})$$

$$= \begin{cases} \bar{a}_i \bullet \bar{x} & \text{if } x \in \Omega \\ d_i & \text{if } x \in \bar{\Omega} \\ \bar{g}_i(\bar{x}, \bar{d}) \leq d_i & \text{otherwise} \end{cases} \quad \text{and } i \text{ is gating} \quad (4)$$

In the above definition, \tilde{g}_i is a function that depends on the problem details. Its region of definition Ψ_i is the intersection of the infeasible region $\bar{\Omega}$ with the complement of the region where component i is the gating component.

[0048] For a production policy such as depicted in FIG. 3, \tilde{g}_i has an additional property: in its region of definition, it is independent of d_i . This is considered the “gate mapping” property. An allocation policy is gate mapping if the following condition holds for every demand vector \bar{x} : If for both $\bar{d} = (d_1, \dots, d_i, \dots, d_n)$ and $\bar{d}' = (d_1, \dots, d'_i, \dots, d_n)$ the gating component is the same (component j , with $j \neq i$), then $\bar{q}(\bar{x}, \bar{d}) = \bar{q}(\bar{x}, \bar{d}')$.

[0049] A corollary is that a production policy is gate mapping if and only if

$$\frac{\partial}{\partial d_i} g_i(\bar{x}, \bar{d}) = 0 \text{ if } \bar{x} \in \Psi_i \quad (5)$$

[0050] For gate mapping production policies, by definition g_i is independent of d_i unless component i is actually gating. This observation relates to the first and third conditions in the second line of Eq. 4. Conversely, if i is gating then $\partial g_i(\bar{x}, \bar{d}) / \partial d_i = 1$ according to the second condition in the second line of Eq. 4. Hence the gating component can easily be identified by examining the derivatives of the g_i .

[0051] The gating risk G_i for component i may be understood as the probability under the distribution of \bar{x} that component i is gating. This probability can be expressed as the expectation:

$$G_i = \int \frac{\partial g_i(\bar{x}, \bar{d})}{\partial d_i} f(\bar{x}, \bar{\mu}, \Sigma) d\bar{x} = \bar{a}_i \cdot \int \frac{\partial \bar{q}(\bar{x}, \bar{d})}{\partial d_i} f(\bar{x}, \bar{\mu}, \Sigma) d\bar{x} \quad (6)$$

[0052] While $\partial \bar{q}(\bar{x}, \bar{d}) / \partial d_i$ may not be differentiable, the above integral may be seen as uniformly convergent in the Lebesgue sense. The order of integration and differentiation can be reversed, to yield the following expression for G_i :

$$G_i = \bar{a}_i \cdot \frac{\partial}{\partial d_i} \int \bar{q}(\bar{x}, \bar{d}) f(\bar{x}, \bar{\mu}, \Sigma) d\bar{x} \quad (7)$$

[0053] The above equation provides the following expression for the gating risk G_i for component i .

$$G_i = \bar{a}_i \cdot \frac{\partial \hat{q}(\bar{d})}{\partial d_i} \quad (8)$$

Where $\hat{q}(\bar{d})$ is the mean production given component availability \bar{d} .

[0054] The mean production $\hat{q}(\bar{d})$ may be understood in comparison to the mean demand value $\bar{\mu}$ 28 discussed above and depicted in FIGS. 2 and 3. In situations where the demand for a set of products falls into an infeasible region of FIG. 2, it is clear that the actual production will not be equal to the demanded production. In such situations, the actual production will be less for one or more products than the actual demand. Because of such situations, the mean production (averaged over the distribution of demand vectors) will be less than the mean demand $\bar{\mu}$ 28 (averaged over the distribution of demand vectors). The mean production may be calculated using multidimensional integrals to average the actual production over the probability distribution of the demanded production. The integration may be performed, for example, using Monte Carlo simulations, quadratures, or other techniques. Additional information on the integration may be found for example, in U.S. Provisional Patent Application No. 60/213,189, referenced above.

[0055] FIG. 4 is a flow diagram illustrating one embodiment of a procedure 40 for calculating gating risk. The procedure 40 commences in block 41 by gathering relevant data concerning components and products. The gathered data include, for example, information such as component consumptions or other information relating to connect rates, information such as component allocations or other information relating to maximum component availabilities, mean product demands or other information relating to expected demands, inter-product correlations or other information relating to the covariance of the relevant demand vector. In block 42, the components to be assessed are selected, for example by a

user input or other input. The procedure then commences looping in block 43 through the various components for which the gating risk is to be determined.

[0056] In block 44a-46, a numerical derivative is calculated of production with respect to the available or expeditable quantity of the component under consideration. As illustrated, a first order difference is used to model the derivative. Other techniques for finding numerical derivatives may be used in addition or instead.

[0057] In block 44a, the component plan is reset to its original value at the maximum expeditable or available plan. The availability of the component under consideration (component i) is then increased by a small increment epsilon. A typical value for epsilon is, for example, 1% of the component plan for component i . In block 44b, the mean production vector q_i^+ is calculated for the increased component plan. The mean production may be calculated using multidimensional integrals over the product space spanned by the demand vector. The integration may be performed, for example, using Monte Carlo simulations, quadratures, or other techniques. Additional information on the integration may be found for example, in U.S. Provisional Patent Application No. 60/213,189, referenced above.

[0058] In block 45a, the component plan is reset to its original value at the maximum expeditable or available plan. The availability of the component under consideration (component i) is then decreased by a small increment epsilon. In block 44b, the mean production vector q_i^- is calculated for the decreased component plan.

[0059] In block 46, the gating risk is calculated through an application of Eq. 8 to the mean productions q_i^+ and q_i^- for the increased and decreased component plan. Block 46 evaluates $a_i * (q_i^+ + q_i^-)/(2 * \text{epsilon})$, where a_i is the vector of connect rates for component i . In block 48 the looping from block 43 ends, and the procedure returns to block 43 to calculate gating risk for the next component (if any). The results of the calculation may be reported to a user in block 49.

[0060] FIG. 5 is a flow diagram illustrating one embodiment of a procedure 50 for calculating shortage risk. In block 51, the relevant data are gathered concerning components and products. In block 52, the components to be assessed are selected, for example by a user

input or other input. The procedure then commences looping in block 53 through the various components for which the shortage risk is to be determined.

[0061] The shortage risk for a given component is evaluated in block 56. The evaluation applies Eq. 2 to calculate the shortage risk r_i for a component i under consideration. In block 58 the looping from block 53 ends, and the procedure returns to block 53 to calculate gating risk for the next component (if any). The results of the calculation may be reported to a user in block 59.

An Example Computing and Network Environment

[0062] ~~Fig. 1~~ FIG. 6 is a block diagram illustrating a network environment in which processing according to embodiments of a system according to the present invention may be practiced. As is illustrated in ~~Fig. 1~~ FIG. 6, network 100, such as a private wide area network (WAN) or the Internet, includes a number of networked servers 110(1)-(N) that are accessible by client computers 120(1)-(N). Communication between client computers 120(1)-(N) and servers 110(1)-(N) typically occurs over a publicly accessible network, such as a public switched telephone network (PSTN), a DSL connection, a cable modem connection or large bandwidth trunks (e.g., communications channels providing T1 or OC3 service). Client computers 120(1)-(N) access servers 110(1)-(N) through, for example, a service provider. This might be, for example, an Internet Service Provider (ISP) such as America On-Line™, Prodigy™, CompuServe™ or the like. Access is typically had by executing application specific software (e.g., network connection software and a browser) on the given one of client computers 120(1)-(N).

[0063] One or more of client computers 120(1)-(N) and/or one or more of servers 110(1)-(N) may be, for example, a computer system of any appropriate design, in general, including a mainframe, a mini-computer or a personal computer system. Such a computer system typically includes a system unit having a system processor and associated volatile and non-volatile memory, one or more display monitors and keyboards, one or more diskette drives, one or more fixed disk storage devices and one or more printers. These computer systems are typically information handling systems which are designed to provide computing power to one or more users, either locally or remotely. Such a computer system may also include one or a plurality of I/O devices (i.e., peripheral devices) which are coupled to the system

processor and which perform specialized functions. Examples of I/O devices include modems, sound and video devices and specialized communication devices. Mass storage devices such as hard disks, CD-ROM drives and magneto-optical drives may also be provided, either as an integrated or peripheral device. One such example computer system, discussed in terms of client computers 120(1)-(N) is shown in detail in ~~Fig. 2~~ FIG. 7.

[0064] It will be noted that the variable identifier "N" is used in several instances in ~~Fig. 1~~ FIG. 6 to more simply designate the final element (e.g., servers 110(1)-(N) and client computers 120(1)-(N)) of a series of related or similar elements (e.g., servers and client computers). The repeated use of such variable identifiers is not meant to imply a correlation between the sizes of such series of elements, although such correlation may exist. The use of such variable identifiers does not require that each series of elements has the same number of elements as another series delimited by the same variable identifier. Rather, in each instance of use, the variable identified by "N" may hold the same or a different value than other instances of the same variable identifier.

[0065] ~~Fig. 2~~ FIG. 7 depicts a block diagram of a computer system 210 suitable for implementing embodiments of the present invention, and ~~example~~ examples of one or more of client computers 120(1)-(N). Computer system 210 includes a bus 212 which interconnects major subsystems of computer system 210 such as a central processor 214, a system memory 216 (typically RAM, but which may also include ROM, flash RAM, or the like), an input/output controller 218, an external audio device such as a speaker system 220 via an audio output interface 222, an external device such as a display screen 224 via display adapter 226, serial ports 228 and 230, a keyboard 232 (interfaced with a keyboard controller 233), a storage interface 234, a floppy disk drive 236 operative to receive a floppy disk 238, and a CD-ROM drive 240 operative to receive a CD-ROM 242. Also included are a mouse 246 (or other point-and-click device, coupled to bus 212 via serial port 228), a modem 247 (coupled to bus 212 via serial port 230) and a network interface 248 (coupled directly to bus 212).

[0066] Bus 212 allows data communication between central processor 214 and system memory 216, which may include both read only memory (ROM) or flash memory (neither shown), and random access memory (RAM) (not shown), as previously noted. The RAM is generally the main memory into which the operating system and application programs are

loaded and typically affords at least 16 megabytes of memory space. The ROM or flash memory may contain, among other code, the Basic Input-Output system (BIOS) which controls basic hardware operation such as the interaction with peripheral components. Applications resident with computer system 210 are generally stored on and accessed via a computer readable medium, such as a hard disk drive (e.g., fixed disk 244), an optical drive (e.g., CD-ROM drive 240), floppy disk unit 236 or other storage medium. Additionally, applications may be in the form of electronic signals modulated in accordance with the application and data communication technology when accessed via network modem 247 or interface 248.

[0067] Storage interface 234, as with the other storage interfaces of computer system 210, may connect to a standard computer readable medium for storage and/or retrieval of information, such as a fixed disk drive 244. Fixed disk drive 244 may be a part of computer system 210 or may be separate and accessed through other interface systems. Many other devices can be connected such as a mouse 246 connected to bus 212 via serial port 228, a modem 247 connected to bus 212 via serial port 230 and a network interface 248 connected directly to bus 212. Modem 247 may provide a direct connection to a remote server via a telephone link or to the Internet via an internet service provider (ISP). Network interface 248 may provide a direct connection to a remote server via a direct network link to the Internet via a POP (point of presence). Network interface 248 may provide such connection using wireless techniques, including digital cellular telephone connection, Cellular Digital Packet Data (CDPD) connection, digital satellite data connection or the like.

[0068] Many other devices or subsystems (not shown) may be connected in a similar manner (e.g., bar code readers, document scanners, digital cameras and so on). Conversely, it is not necessary for all of the devices shown in ~~Fig. 2~~ FIG. 7 to be present to practice the present invention. The devices and subsystems may be interconnected in different ways from that shown in ~~Fig. 2~~ FIG. 7. The operation of a computer system such as that shown in ~~Fig. 2~~ FIG. 7 is readily known in the art and is not discussed in detail in this application. Code to implement the present invention may be stored in computer-readable storage media such as one or more of system memory 216, fixed disk 244, CD-ROM 242, or floppy disk 238. Additionally, computer system 210 may be any kind of computing device, and so includes personal data assistants (PDAs), network appliance, X-window terminal or other such computing device. The operating system provided on computer system 210 may be MS-

DOS®, MS-WINDOWS®, OS/2®, UNIX®, Linux® or other known operating system. Computer system 210 also supports a number of Internet access tools, including, for example, an HTTP-compliant web browser having a JavaScript interpreter, such as Netscape Navigator® 3.0, Microsoft Explorer® 3.0 and the like.

[0069] Moreover, regarding the signals described herein, those skilled in the art will recognize that a signal may be directly transmitted from a first block to a second block, or a signal may be modified (e.g., amplified, attenuated, delayed, latched, buffered, inverted, filtered or otherwise modified) between the blocks. Although the signals of the above described embodiment are characterized as transmitted from one block to the next, other embodiments of the present invention may include modified signals in place of such directly transmitted signals as long as the informational and/or functional aspect of the signal is transmitted between blocks. To some extent, a signal input at a second block may be conceptualized as a second signal derived from a first signal output from a first block due to physical limitations of the circuitry involved (e.g., there will inevitably be some attenuation and delay). Therefore, as used herein, a second signal derived from a first signal includes the first signal or any modifications to the first signal, whether due to circuit limitations or due to passage through other circuit elements which do not change the informational and/or final functional aspect of the first signal.

[0070] The foregoing described embodiment wherein the different components are contained within different other components (e.g., the various elements shown as components of computer system 210). It is to be understood that such depicted architectures are merely examples, and that in fact many other architectures can be implemented which achieve the same functionality. In an abstract, but still definite sense, any arrangement of components to achieve the same functionality is effectively “associated” such that the desired functionality is achieved. Hence, any two components herein combined to achieve a particular functionality can be seen as “associated with” each other such that the desired functionality is achieved, irrespective of architectures or intermediate components. Likewise, any two components so associated can also be viewed as being “operably connected”, or “operably coupled”, to each other to achieve the desired functionality.

[0071] ~~Fig. 3~~ FIG. 8 is a block diagram depicting a network 300 in which computer system 210 is coupled to an internetwork 310, which is coupled, in turn, to client systems 320.

and 330, as well as a server 340. Internetwork 310 (e.g., the Internet) is also capable of coupling client systems 320 and 330, and server 340 to one another. With reference to computer system 210, modem 247, network interface 248 or some other method can be used to provide connectivity from computer system 210 to internetwork 310. Computer system 210, client system 320 and client system 330 are able to access information on server 340 using, for example, a web browser (not shown). Such a web browser allows computer system 210, as well as client systems 320 and 330, to access data on server 340 representing the pages of a website hosted on server 340. Protocols for exchanging data via the Internet are well known to those skilled in the art. Although Fig. 3 FIG. 8 depicts the use of the Internet for exchanging data, the present invention is not limited to the Internet or any particular network-based environment.

[0072] Referring to Figs. 1, 2 and 3, a browser running on computer system 210 employs a TCP/IP connection to pass a request to server 340, which can run an HTTP “service” (e.g., under the WINDOWS® operating system) or a “daemon” (e.g., under the UNIX® operating system), for example. Such a request can be [[processed,]] processed, for example, by contacting an HTTP server employing a protocol that can be used to communicate between the HTTP server and the client computer. The HTTP server then responds to the protocol, typically by sending a “web page” formatted as an HTML file. The browser interprets the HTML file and may form a visual representation of the same using local resources (e.g., fonts and colors).

An Example Architecture for Estimation of Component Gating and Shortage Risk in Manufacturing Operations

[0073] — Please refer to Appendix A, which is attached hereto and is incorporated herein by reference, in its entirety and for all purposes. Moreover, in addition to the description of the method of the present invention contained within Appendix A, further information is available in the patent applications now incorporated by reference herein: This invention is related to Provisional Application No. 60/213,189, filed 6/21/2000, which is incorporated by reference herein. It is also related to U.S. Patent Application Nos. 09/412,560, filed 10/5/1999, and 09/491,461, filed 1/26/2000, both of which are incorporated by reference herein.

[0074] Various processes according to embodiments of the present invention are discussed herein. It is appreciated that operations discussed herein may consist of directly entered commands by a computer system user or by steps executed by application specific hardware modules, but the preferred embodiment includes steps executed by software modules. The functionality of steps referred to herein may correspond to the functionality of modules or portions of modules.

[0075] The operations referred to herein may be modules or portions of modules (e.g., software, firmware or hardware modules). For example, although the described embodiment includes software modules and/or includes manually entered user commands, the various example modules may be application specific hardware modules. The software modules discussed herein may include script, batch or other executable files, or combinations and/or portions of such files. The software modules may include a computer program or subroutines thereof encoded on computer-readable media.

[0076] Additionally, those skilled in the art will recognize that the boundaries between modules are merely illustrative and alternative embodiments may merge modules or impose an alternative decomposition of functionality of modules. For example, the modules discussed herein may be decomposed into submodules to be executed as multiple computer processes, and, optionally, on multiple computers. Moreover, alternative embodiments may combine multiple instances of a particular module or submodule. Furthermore, those skilled in the art will recognize that the operations described in example embodiment are for illustration only. Operations may be combined or the functionality of the operations may be distributed in additional operations in accordance with the invention.

[0077] Alternatively, such actions may be embodied in the structure of circuitry that implements such functionality, such as the micro-code of a complex instruction set computer (CISC), firmware programmed into programmable or erasable/programmable devices, the configuration of a field-programmable gate array (FPGA), the design of a gate array or full-custom application-specific integrated circuit (ASIC), or the like.

[0078] Each of the blocks of the flow diagram may be executed by a module (e.g., a software module) or a portion of a module or a computer system user using, for example, a computer system such as the storage router previously mentioned, or a similar network

element, as well as a computer system such as computer system 210. Thus, the above described method, the operations thereof and modules ~~[[therefor]]~~ thereof may be executed on a computer system configured to execute the operations of the method and/or may be executed from computer-readable media. The method may be embodied in a machine-readable and/or computer-readable medium for configuring a computer system to execute the method. Thus, the software modules may be stored within and/or transmitted to a computer system memory to configure the computer system to perform the functions of the module.

[0079] Such a computer system normally processes information according to a program (a list of internally stored instructions such as a particular application program and/or an operating system) and produces resultant output information via I/O devices. A computer process typically includes an executing (running) program or portion of a program, current program values and state information, and the resources used by the operating system to manage the execution of the process. A parent process may spawn other, child processes to help perform the overall functionality of the parent process. Because the parent process specifically spawns the child processes to perform a portion of the overall functionality of the parent process, the functions performed by child processes (and grandchild processes, etc.) may sometimes be described as being performed by the parent process.

[0080] Such a computer system typically includes multiple computer processes executing “concurrently.” Often, a computer system includes a single processing unit which is capable of supporting many active processes alternately. Although multiple processes may appear to be executing concurrently, at any given point in time only one process is actually executed by the single processing unit. By rapidly changing the process executing, a computer system gives the appearance of concurrent process execution. The ability of a computer system to multiplex the computer system’s resources among multiple processes in various stages of execution is called multitasking. Systems with multiple processing units, which by definition can support true concurrent processing, are called multiprocessing systems. Active processes are often referred to as executing concurrently when such processes are executed in a multitasking and/or a multiprocessing environment.

[0081] The software modules described herein may be received by such a computer system, for example, from computer readable media. The computer readable media may be permanently, removably or remotely coupled to the computer system. The computer readable

media may non-exclusively include, for example, any number of the following: magnetic storage media including disk and tape storage media; optical storage media such as compact disk media (e.g., CD-ROM, CD-R, etc.) and digital video disk storage media; nonvolatile memory storage memory including semiconductor-based memory units such as FLASH memory, EEPROM, EPROM, ROM or application specific integrated circuits, volatile storage media including registers, buffers or caches, main memory, RAM, and the like; and data transmission media including computer network, point-to-point telecommunication, and carrier wave transmission media. In a UNIX-based embodiment, the software modules may be embodied in a file which may be a device, a terminal, a local or remote file, a socket, a network connection, a signal, or other expedient of communication or state change. Other new and various types of computer-readable media may be used to store and/or transmit the software modules discussed herein.

[0082] While particular embodiments of the present invention have been shown and described, it will be obvious to those skilled in the art that, based upon the teachings herein, changes and modifications may be made without departing from this invention and its broader aspects and, therefore, the appended claims are to encompass within their scope all such changes and modifications as are within the true spirit and scope of this invention. Furthermore, it is to be understood that the invention is solely defined by the appended claims.

METHOD AND APPARATUS FOR ESTIMATION OF COMPONENT GATING AND SHORTAGE RISK IN MANUFACTURING OPERATIONS

Balazs Kralik, Michael Goldbach, and Paul Dagum

ABSTRACT OF THE DISCLOSURE

A method for estimation of component gating risk in manufacturing operations is disclosed. The method includes generating an altered component plan by altering a component plan for a component, computing a mean production value using the altered component plan, and computing the component gating risk using the mean production value.

~~APPENDIX A~~

Provisional Patent Application of
5 Balazs Kralik, Michael Goldbach, and Paul Dagum
for
Method and Business Process for Estimation of Component Gating
and Shortage Risks in Assemble-to-Order Manufacturing Operations

10 CROSS-REFERENCE TO RELATED APPLICATIONS

This invention is related to Provisional Application No.
60/213,189, filed 6/21/2000, which is herein incorporated by
reference. It is also related to U.S. Patent Application Nos.
09/412,560, filed 10/5/1999, and 09/491,461, filed 1/26/2000,
15 both of which are herein incorporated by reference.

FIELD OF THE INVENTION

This invention relates generally to manufacturing resource
allocation. More particularly, it relates to risk management in
20 the procurement of unfinished goods by estimation of component
shortage risk.

SUMMARY

An assemble-to-order (ATO) manufacturing process is one in which
25 products are manufactured from raw components only as orders are
received. While ATO processes are more efficient than
manufacture-to-stock (MTS) operations, in which products are
produced before demand is known, they are not without
significant risks. If component availability is below what is
30 required to meet product demand, then fulfillment of orders is
either delayed or fails. Conversely, if excess component
inventory remains after demand is met, the excess is either
returned to the supplier, usually at a loss, or held until
future need. Thus any component inventory that does not exactly

meet demand incurs some financial loss for the manufacturer. Being able to predict the risk of component shortage for each component allows a manufacturer to make more informed business decisions.

5

The present invention addresses this need by providing a novel method for estimating component shortage risk and component gating risk. A gating component is one that is the most short and therefore determines the allowable level of production.

10

Gating risk for a particular component is the risk that the component will be the gating component. Estimation of component shortage and gating risks is particularly complicated when a large number of products are made from a large number of the same components, and thus the present invention is particularly advantageous under such conditions.

15

Full description of the method of the present invention is contained within the following attached appendices as well as in the patent applications incorporated by reference above:

20

Appendix A: Invention Disclosure ID-07-00, "Method and business process for the estimation of component gating and shortage risk in assemble-to-order manufacturing operations," Balazs Kralik et al.

25

Appendix B: Technical Report TR-26-00, "Gating and Shortage Risk," Balazs Kralik and Michael Goldbach.

30

In addition to a business process and computer-implemented method, the present invention provides a system for computing component gating and shortage risks. The system is preferably a computer system containing a memory, input/output devices, and processing means for implementing the method.

The present invention also provides a computer-readable medium of instructions executable by the computer to perform method steps for the method described above. The present invention applies equally regardless of the particular type of signal
5 bearing media used. Examples of computer-readable media include: recordable-type media such as floppy disks and CD-ROMs and transmission-type media such as digital and analog communication links.

10 It will be clear to one skilled in the art that the above embodiment may be altered in many ways without departing from the scope of the invention.

~~rap~~t

Appendix A:
Invention Disclosure
ID-07-00

STRICTLY CONFIDENTIAL

Title: Method and business process for the estimation of component gating and shortage risk in assemble-to-order manufacturing operations

Inventor(s): Balazs Kralik, Michael Goldbach, Paul Dagum

<u>Background of the invention</u>	3
<u>Problem definition</u>	3
<u>Indication of Novelty</u>	3
<u>Summary of the Invention</u>	3
<u>Detailed Description of a Preferred Embodiment</u>	4
<u>Technical field of the invention</u>	4
<u>Statutory Classification of the invention:</u>	7
<u>Inventors information:</u>	7

Background of the invention

Problem definition

When possible, it is beneficial for manufacturing companies to produce their outputs in an assemble-to-order (ATO) fashion. While manufacture-to-stock (MTS) manufacturers produce outputs before demand for them is revealed, ATO operations only produce outputs as the order stream arrives.

In an ATO operation, the availability of components is the key variable that controls the ability to produce. If component availability is limited (either because inventory is limited or because supplier commitments to ship the component on demand are insufficient), then fulfillment of demand either fails or is delayed.

If component planners must negotiate component supply contracts (specifying the number of each component to ship in the planning period) before demand is realized, there is a residual risk that some components will be in short supply. Estimating this risk is paramount for informed supply contract negotiation.

Component Shortage Risk quantifies the risk that in the planning period there will be more need than availability for a given component. In other words Component Shortage Risk for component i is the risk that there will be a shortage in component i .

Component Gating Risk quantifies the risk that a particular component will be the cause of order fulfillment failure. Note that in many demand instances more than one component may be short, but one of the components, the one that is most short, is the *gating component*, i.e. the component determining the level of production. Component Gating Risk for component i is the risk that component i will be the gating component.

Problems Solved by the Invention:

Problem 1: How to compute the value of component gating risk

Problem 2: How to compute the value of component shortage risk

This invention makes use of the method disclosed in [1], which in turn makes use of the method disclosed in [2].

Indication of Novelty

The rigorous definition and method of computing gating risk is novel.

Both gating risk and shortage risk computations represents novel uses of the invention disclosed in [1].

Summary of the Invention

The risk measures are computed following these steps
Step 1. Gather the relevant component and product data

- Step 2. Indicate the components for which to compute gating and/or shortage risk
 Step 3. For all components indicated in step 2, compute gating and/or shortage risk.
 Step 4. Return the results

Step 3. contains the novelty and is expanded below.

Detailed Description of a Preferred Embodiment

Gating risk

The first part of the present invention, illustrated in Figure 1, describes a means to compute the gating risk for a set of components.

In step 101, we gather the component and product data. The required data are exactly the ones described in [1].

In step 102, the user specifies which component for which the gating risk should be computed.

In step 103-110 we loop over the components for which gating risk computation is desired.

In steps 105 through 109, we compute the numerical derivative of production with respect to the plan for each component i . In the preferred embodiment, a first order difference is used, however other numerical derivatives could be applied.

In step 105 we first reset the component plan to its original value at the maximum expeditable plan, and then increase the component plan of component i under consideration by a small amount epsilon. Epsilon is set typically to 1% of the component plan of component i .

In step 106 we use the method disclosed in [1] to compute q_{i+} , the mean production under increased component plan i .

In step 107 we first reset the component plan to its original value at the maximum expeditable plan, and then decrease the component plan of component i under consideration by a small amount epsilon.

In step 108 we use the method disclosed in [1] to compute q_{i-} , the mean production under decreased component plan i .

In step 109 we apply formula 3.6 of [3] to compute the gating risk. Explicitly, we evaluate $a_i * (q_{i+} - q_{i-}) / (2 * \epsilon)$ where a_i is the vector of connect rates for component i and q_{i+} , q_{i-} and ϵ have been defined above.

In step 111 the results of the gating risk computation are returned to the user.

Shortage Risk

The second part of the invention is the computation of shortage risk. This consists of the steps described in Figure 2:

In step 201 the component and product parameters are collected. The required parameters are exactly the same as those required for the method described in [1].

In step 202 the user specifies which components the shortage risk is to be calculated for.

In step 203-205 we loop over all the selected components

In step 204 we perform the calculation according to formula 2.1 of the attached technical report, TR-26-00.

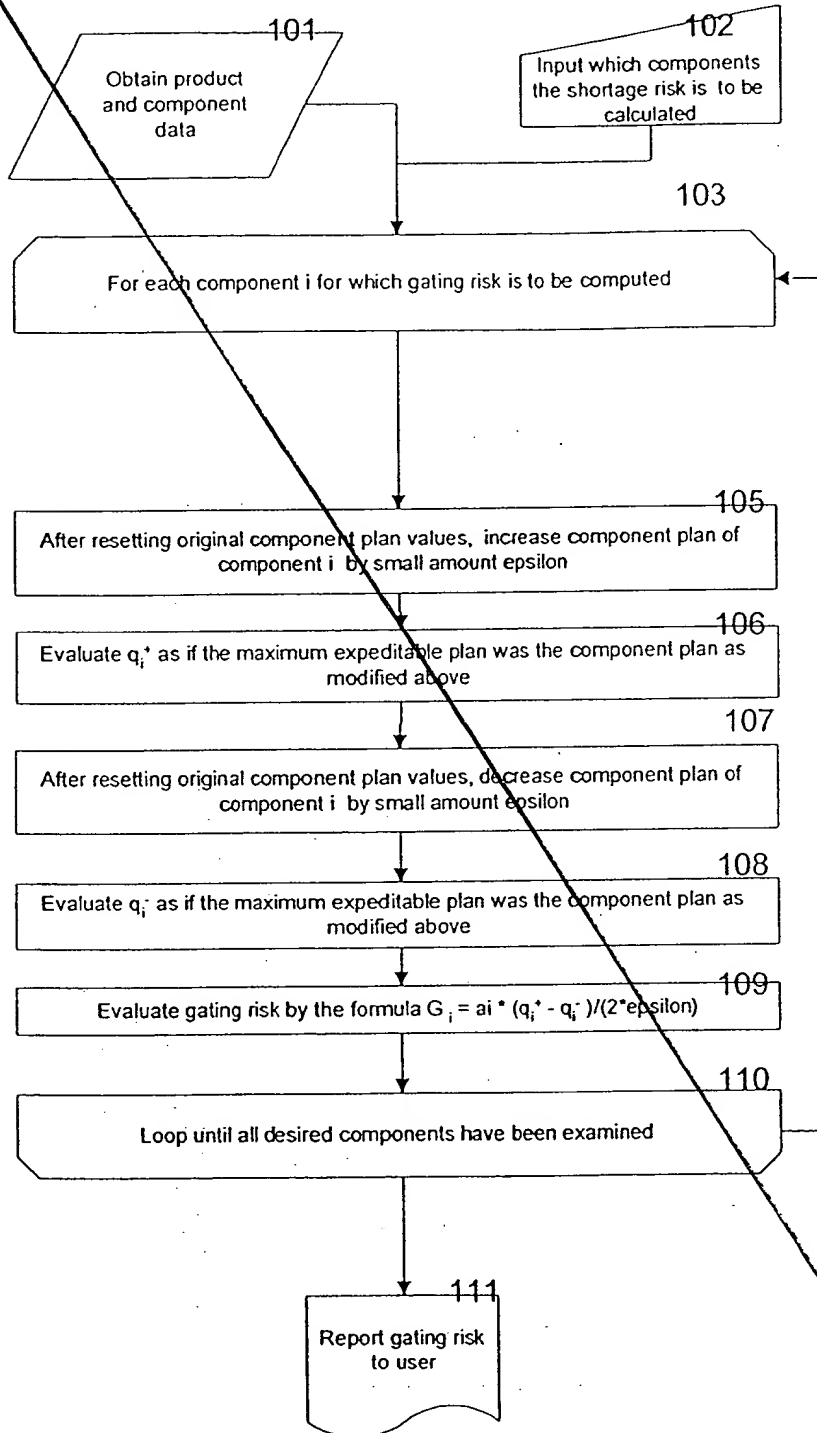


Figure 1 Computing gating risk

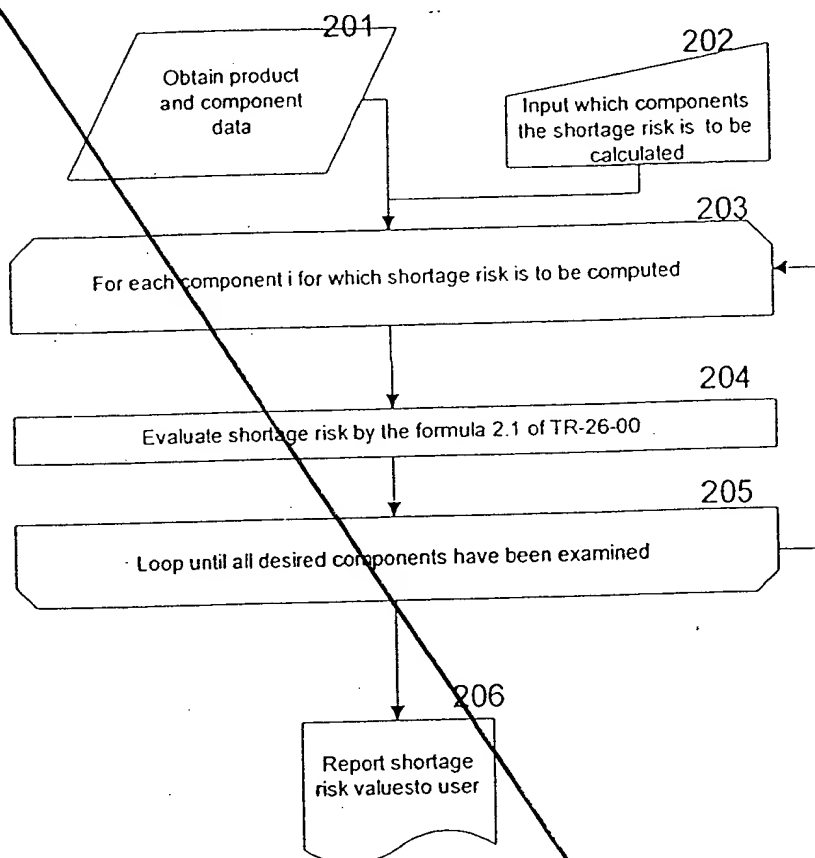


Figure 2 Computing shortage risk

Technical field of the invention

This invention relates generally to manufacturing resource planning. More specifically it relates to risk management in the procurement of unfinished goods.

Statutory Classification of the invention:

- This invention represents a system and method

Inventors information:

Full legal name, including middle name, if any: **Balazs Kralik**
 Residential address: **2926 Hillegass Ave, Berkeley CA 94705**
 Postal address:
 Work address: **324 Ritch Street, San Francisco CA 94107**
 Citizenship: **Hungary**
 Home Telephone Number: **510-845-7681**
 Work Telephone Number: **415-284-5815**
 Fax Number: **415-247-9768**
 E-Mail address: **balazs.kralik@rapt.com**

Full legal name, including middle name, if any: **Michael Goldbach**
 Residential address: **1004 Emerson Street, Palo Alto CA 94301**

Postal address:

Work address: 324 Ritch Street, San Francisco CA 94107

Citizenship: Canada

Home Telephone Number: 650-325-9810

Work Telephone Number: 415-284-5807

Fax Number: 415 -247-9768

E-Mail address: michael.goldbach@rapt.com

Full legal name, including middle name, if any: Paul Dagum

Residential address: 435 Locust Street, San Francisco CA 94018

Postal address:

Work address: 324 Ritch Street, San Francisco CA 94107

Citizenship: Canada

Home Telephone Number: 415-409-1109

Work Telephone Number: 415-284-5802

Fax Number: 415 -247-9768

E-Mail address: paul.dagum@rapt.com

References

1. Dagum, P., M. Goldbach, and B. Kralik, *Method and apparatus for the estimation of mean production for assemble-to-order manufacturing operations*. U.S. Provisional Patent Application No. 60/213,189, filed 6/21/2000.
2. Chavez, T.A. and P. Dagum, *Method and Apparatus for Multivariate Allocation of Resources*. Rapt Inc., 1999. US Patent Application No. 09/412560, filed 10/5/1999.
3. Kralik, B. and M. Goldbach, *Gating and Shortage Risk*. 2000: Rapt Inc. TR-26-00

Appendix B

Gating and Shortage Risk

Balazs Kralik and Michael Goldbach

Technical Report TR-26-00

07-jul-00

PROPRIETARY AND CONFIDENTIAL
DO NOT DISTRIBUTE

Contents

1	Introduction	3
2	Shortage Risk	3
3	Gating Risk	5
3.1	Evaluating Gating Risk	6

List of Figures

1	The 2-product 2-component (p2c2) model	4
2	Regions of shortage for component 1, component 2 and both for the p2c2 model . . .	4
3	Regions where either component 1 or component 2 is gating for the p2c2 model . . .	6

1 Introduction

When possible, it is beneficial for manufacturing companies to produce their outputs in an assemble-to-order (ATO) fashion. While manufacture-to-stock (MTS) manufacturers produce outputs before demand for them is revealed, ATO operations only produce outputs as the order stream arrives.

In an ATO operation, the availability of components is the key variable that controls the ability to produce. If component availability is limited (either because inventory is limited or because supplier commitments to ship the component on demand are insufficient), then fulfillment of demand either fails or is delayed.

If component planners must negotiate component supply contracts (specifying the number of each component to ship in the planning period) before demand is realized, there is a residual risk that some components will be in shortage (unless the supply contract calls for the delivery of infinitely many components). Estimating this risk is paramount for informed supply contract negotiation.

Component Shortage Risk quantifies the risk that in the planning period there will be more need than availability for this component. In other words Component Shortage Risk for component i is the risk that there will be a shortage in component i .

Component Gating Risk quantifies the risk that a particular component will be the most important cause for order fulfillment failure. Note that in many demand instances more than one component may be short, but one of the components, the one that is shortest, is the gating component, i.e. the component determining the level of production. Component Gating Risk for component i is the risk that component i will be the gating component.

2 Shortage Risk

Production is feasible if demand x falls in the region $\Omega = \{x : a_i \cdot x \leq d_i \ \forall i\}$, where a_i are the connect rates and d_i is the maximum component availability for component i . In practical terms, d_i are the *maximum expeditable* levels.

A shortage event occurs if the realized demand x is such that $a_i \cdot x > d_i$. To illustrate the concept, consider the simple case of the p2c2 model illustrated in Figure 1. In this model the regions of demand which correspond to shortage are illustrated in Figure 2.

The shortage risk is always given by the weight of the normal distribution over a halfspace.

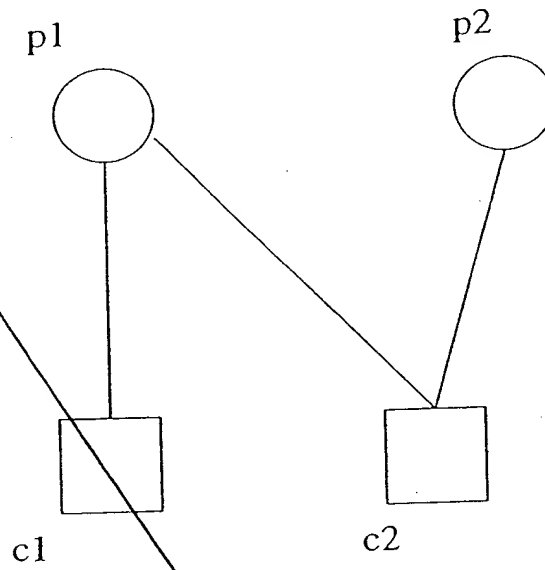


Figure 1: The 2-product 2-component (p2c2) model

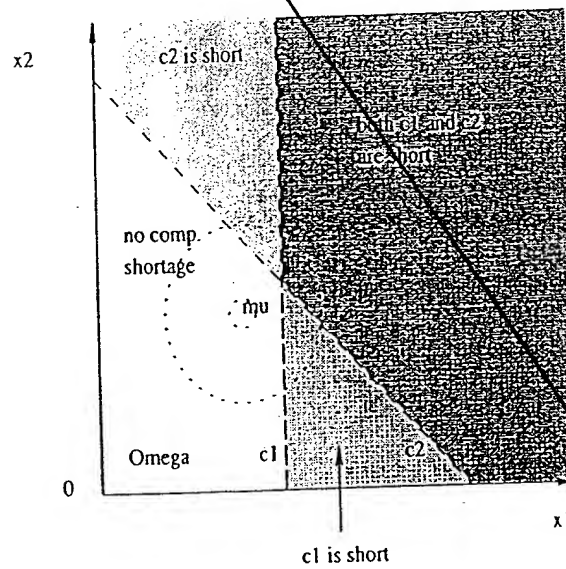


Figure 2: Regions of shortage for component 1, component 2 and both for the p2c2 model

The shortage risk r_i for component i is given by

$$r_i = \int_{d_i}^{\infty} N\left(x; a_i \cdot \mu, \sqrt{a_i^T \Sigma a_i}\right) dx \quad (2.1)$$

where d_i is the maximum expeditable level of component i , $N(x, \mu, \sigma)$ is the normal density with mean μ and variance σ^2 , a_i is the vector of connect rates for component i , μ is the mean (forecasted) demand, and Σ represents demand covariance.

3 Gating Risk

The regions of gating risk are illustrated in Figure 3. Intuitively, we call component i the gating component given demand x and component availability d if the production $q(x, d)$ falls on the d_i -face of the feasible region Ω .

Definition 3.1 Given a demand event x , a component plan d , and a production policy $q : x \in \mathbb{R}^n, d \in \mathbb{R}^m \rightarrow \mathbb{R}^n$, we call component i gating if and only if $q(x, d)$ falls on the face of the feasible region Ω associated with component i . In other words if and only if

$$a_i \cdot q(x, d) = d_i. \quad (3.1)$$

Note: For the local-u production policy (depicted in Figure 3) it is easy to see that the space where more than one component is gating is of measure zero, except in the degenerate case where there are two components i_1 and i_2 such that $a_{i_1} = a_{i_2}$ and $d_{i_1} = d_{i_2}$. Exclude this degenerate case from the following discussion.

We can paraphrase the definition of gating risk as: given a demand x , whichever component actually runs out completely upon producing the production amount $q(x, d)$ is the gating component.

In order to proceed, consider the function

$$g_i(x, d) = a_i \cdot q(x, d) \quad (3.2)$$

$$= \begin{cases} a_i \cdot x & \text{if } x \in \Omega \\ d_i & \text{if } x \in \bar{\Omega} \text{ and } i \text{ is gating} \\ \bar{g}_i(x, d) \leq d_i & \text{otherwise.} \end{cases} \quad (3.3)$$

In the above definition, \bar{g}_i is some function which depends on the problem details. Its region of definition Ψ_i is the intersection of $\bar{\Omega}$ with the complement of the region where i is gating.

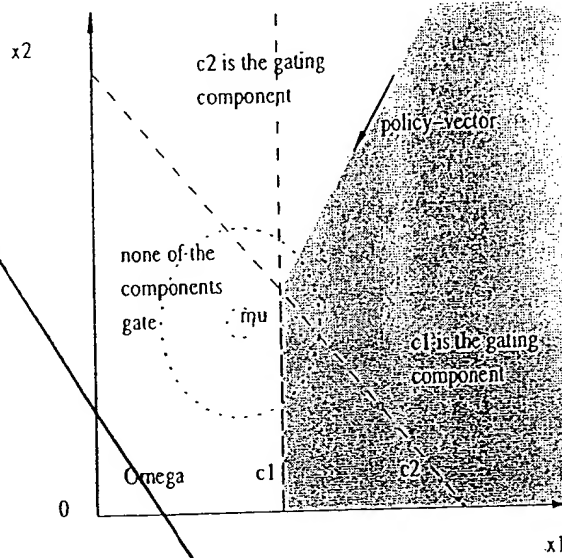


Figure 3: Regions where either component 1 or component 2 is gating for the p2c2 model

For the usual local- u production policy [1], \bar{g}_i has an additional property: in its region of definition, it is independent of d_i . This is the *gate mapping property*:

Definition 3.2 *Gate mapping property*: an allocation policy is gate mapping if the following condition holds for every x : If for both $d = (d_1, \dots, d_i, \dots, d_n)$ and $d' = (d_1, \dots, d'_i, \dots, d_n)$ the gating component is the same $j \neq i$, then $q(x, d) = q(x, d')$.

Corollary 3.3 A production policy is gate mapping if and only if

$$\frac{\partial g_i(x, d)}{\partial d_i} = 0 \quad \text{if } x \in \Psi_i \quad (3.4)$$

3.1 Evaluating Gating Risk

For gate mapping production policies, by definition g_i is independent of d_i unless component i is actually gating. Conversely, if i is gating then $\partial g_i(x, d) / \partial d_i = 1$ by Equation 3.2. Hence the gating component can easily be identified by examining the derivatives of the g_i .

The gating risk G_i for component i is defined as the probability under the distribution of x that component i is gating. This can be expressed as the expectation:

$$G_i = \int \frac{\partial g_i(x, d)}{\partial d_i} f(x; \mu, \Sigma) dx = a_i \cdot \int \frac{\partial q(x, d)}{\partial d_i} f(x; \mu, \Sigma) dx \quad (3.5)$$

While $\frac{\partial q(x, d)}{\partial d_i}$ is not differentiable, the integral is uniformly convergent in the Lebesgue sense, and hence the order of integration and differentiation can be reversed:

$$G_i = a_i \cdot \frac{\partial}{\partial d_i} \int q(x, d) f(x; \mu, \Sigma) dx = a_i \cdot \frac{\partial \bar{q}(d)}{\partial d_i} \quad (3.6)$$

where $\bar{q}(d)$ is the mean production given component availability d .

References

- [1] Paul Dagum, Michael Goldbach, and Balazs Kralik. Demand statistical arbitrage with variants of the uniform production policy. Technical Report TR-14-00, Rapt Technologies Corporation, 324 Ritch Street, San Francisco, CA 94107, January 2000.

~~Script~~

Invention Disclosure

ID-07-00

STRICTLY CONFIDENTIAL

Title: Method and business process for the estimation of component gating and shortage risk in assemble-to-order manufacturing operations

Inventor(s): Balazs Kralik, Michael Goldbach, Paul Dagum

<u>Background of the invention</u>	3
<u>Problem definition</u>	3
<u>Description of the Prior Art</u>	3
<u>Indication of Novelty</u>	3
<u>Indication of Unobviousness</u>	3
<u>Summary of the Invention</u>	4
<u>Detailed Description of a Preferred Embodiment</u>	4
<u>Technical field of the invention</u>	4
<u>Statutory Classification of the invention:</u>	7
<u>Conception and Reduction to Practice</u>	7
<u>Public Disclosures and Statutory Bars</u>	7
<u>Prior Art (references)</u>	7
<u>Inventors information:</u>	8
<u>Statement and Signature(s)</u>	8
<u>Inventor(s)</u>	8
<u>Witness</u>	9

Background of the invention

Problem definition

When possible, it is beneficial for manufacturing companies to produce their outputs in an assemble-to-order (ATO) fashion. While manufacture-to-stock (MTS) manufacturers produce outputs before demand for them is revealed, ATO operations only produce outputs as the order stream arrives.

In an ATO operation, the availability of components is the key variable that controls the ability to produce. If component availability is limited (either because inventory is limited or because supplier commitments to ship the component on demand are insufficient), then fulfillment of demand either fails or is delayed.

If component planners must negotiate component supply contracts (specifying the number of each component to ship in the planning period) before demand is realized, there is a residual risk that some components will be in short supply. Estimating this risk is paramount for informed supply contract negotiation.

Component Shortage Risk quantifies the risk that in the planning period there will be more need than availability for a given component. In other words Component Shortage Risk for component i is the risk that there will be a shortage in component i .

Component Gating Risk quantifies the risk that a particular component will be the cause of order fulfillment failure. Note that in many demand instances more than one component may be short, but one of the components, the one that is most short, is the *gating component*, i.e. the component determining the level of production. Component Gating Risk for component i is the risk that component i will be the gating component.

Problems Solved by the Invention:

Problem 1: How to compute the value of component gating risk

Problem 2: How to compute the value of component shortage risk

Description of the Prior Art

This invention makes use of the method disclosed in [1], which in turn makes use of the method disclosed in [2].

Indication of Novelty

The rigorous definition and method of computing gating risk is novel.

Both gating risk and shortage risk computations represents novel uses of the invention disclosed in [1].

Indication of Unobviousness

Summary of the Invention

The risk measures are computed following these steps

Step 1. Gather the relevant component and product data

Step 2. Indicate the components for which to compute gating and/or shortage risk

Step 3. For all components indicated in step 2, compute gating and/or shortage risk.

Step 4. Return the results

Step 3. contains the novelty and is expanded below.

Detailed Description of a Preferred Embodiment

Gating risk

The first part of the present invention, illustrated in Figure 1, describes a means to compute the gating risk for a set of components.

In step 101, we gather the component and product data. The required data are exactly the ones described in [1].

In step 102, the user specifies which component for which the gating risk should be computed.

In step 103-110 we loop over the components for which gating risk computation is desired.

In steps 105 through 109, we compute the numerical derivative of production with respect to the plan for each component i . In the preferred embodiment, a first order difference is used, however other numerical derivatives could be applied.

In step 105 we first reset the component plan to its original value at the maximum expeditable plan, and then increase the component plan of component i under consideration by a small amount epsilon. Epsilon is set typically to 1% of the component plan of component i .

In step 106 we use the method disclosed in [1] to compute q_{i+} , the mean production under increased component plan i .

In step 107 we first reset the component plan to its original value at the maximum expeditable plan, and then decrease the component plan of component i under consideration by a small amount epsilon.

In step 108 we use the method disclosed in [1] to compute q_{i-} , the mean production under decreased component plan i .

In step 109 we apply formula 3.6 of [3] to compute the gating risk. Explicitly, we evaluate $a_i * (q_{i+} - q_{i-}) / (2 * \epsilon)$ where a_i is the vector of connect rates for component i and q_{i+} , q_{i-} and epsilon have been defined above.

In step 111 the results of the gating risk computation are returned to the user.

Shortage Risk

The second part of the invention is the computation of shortage risk. This consists of the steps described in Figure 2:

In step 201 the component and product parameters are collected. The required parameters are exactly the same as those required for the method described in [1].

In step 202 the user specifies which components the shortage risk is to be calculated for.

In step 203-205 we loop over all the selected components

In step 204 we perform the calculation according to formula 2.1 of the attached technical report.

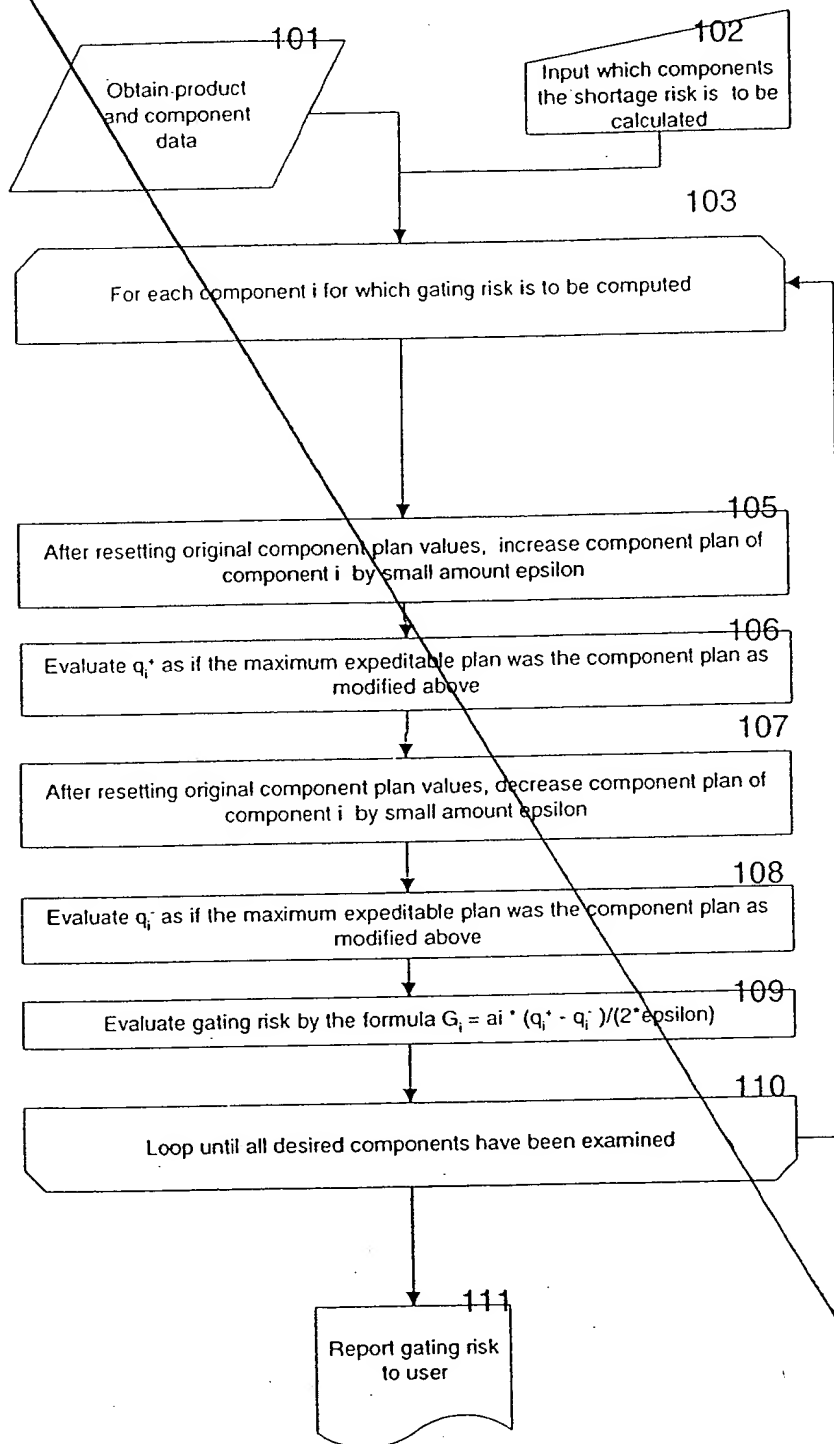


Figure 1 Computing gating risk

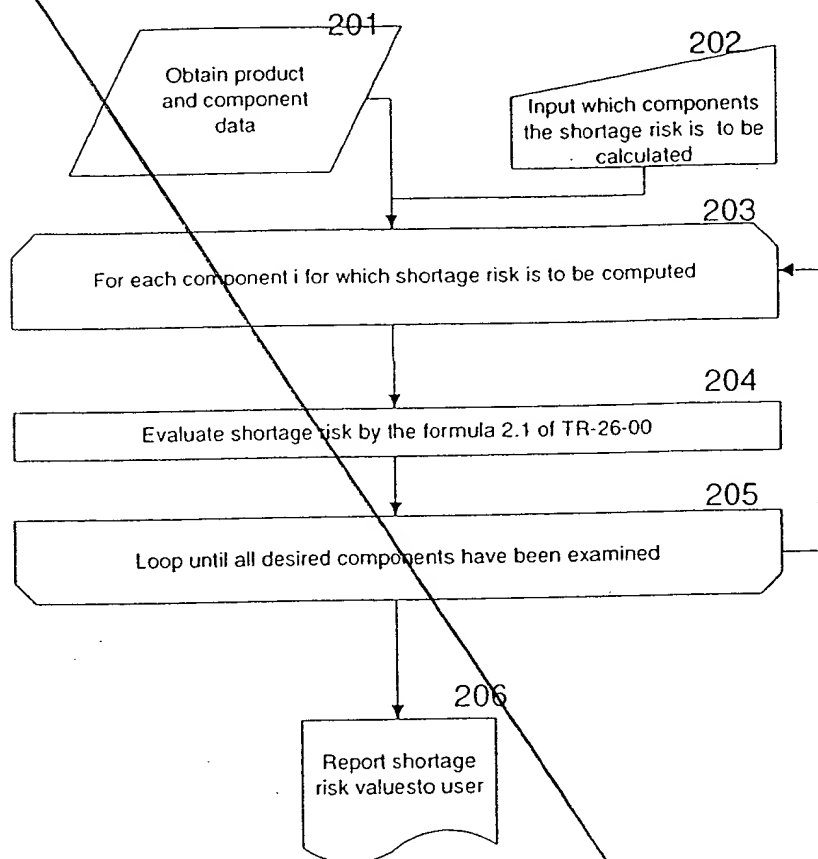


Figure 2 Computing shortage risk

Technical field of the invention

This invention relates generally to manufacturing resource planning. More specifically it relates to risk management in the procurement of unfinished goods.

Statutory Classification of the invention:

- This invention represents a system and method

Conception and Reduction to Practice

- Date of conception: January 2000
- Date of first prototype: February 2000

Public Disclosures and Statutory Bars

- Computation of gating risk was included in Rapt Buy released in March 2000.

Prior Art

Inventors information:

Full legal name, including middle name, if any: **Balazs Kralik**
Residential address: **2926 Hillegass Ave, Berkeley CA 94705**
Postal address:
Work address: **324 Ritch Street, San Francisco CA 94107**
Citizenship: **Hungary**
Home Telephone Number: **510-845-7681**
Work Telephone Number: **415-284-5815**
Fax Number: **415 -247-9768**
E-Mail address: **balazs.kralik@rapt.com**

Full legal name, including middle name, if any: **Michael Goldbach**
Residential address: **1004 Emerson Street, Palo Alto CA 94301**
Postal address:
Work address: **324 Ritch Street, San Francisco CA 94107**
Citizenship: **Canada**
Home Telephone Number: **650-325-9810**
Work Telephone Number: **415-284-5807**
Fax Number: **415 -247-9768**
E-Mail address: **michael.goldbach@rapt.com**

Full legal name, including middle name, if any: **Paul Dagum**
Residential address: **435 Locust Street, San Francisco CA 94018**
Postal address:
Work address: **324 Ritch Street, San Francisco CA 94107**
Citizenship: **Canada**
Home Telephone Number: **415-409-1109**
Work Telephone Number: **415-284-5802**
Fax Number: **415 -247-9768**
E-Mail address: **paul.dagum@rapt.com**

Statement and Signature(s)***Inventor(s)***

As an undersigned inventor, I declare that I have reviewed and understood the contents of the above information and description, and that all the information and description provided above in relation to this invention is correct and complete.

By: _____

Name: _____

Date: _____

By: _____

Name: _____

Date: _____

Witness

I declare that I have read and understood the entire contents of the above invention disclosure on or before the date of my signature below.

By: _____

Name: _____

Date: _____

References

1. Dagum, P., M. Goldbach, and B. Kralik, *Method and business process for the estimation of mean production for assemble-to-order manufacturing operations*. Rapt Inc., 2000. RAPT ID-04-00.
2. Chavez, T.A. and P. Dagum, *Method and Apparatus for Multivariate Allocation of Resources*. Rapt Inc., 1999. US 09/412560.
3. Kralik, B. and M. Goldbach, *Gating and Shortage Risk*. 2000: Rapt Inc. TR-26-00